

VOICE RECOGNITION DEVICE AND METHOD OF VOICE DATA RECOGNITION

Publication number: JP2000259172

Publication date: 2000-09-22

Inventor: SUZUKI TAKASHI

Applicant: CANON KK

Classification:

- international: G10L15/10; G10L15/06; G10L15/22; G10L15/24;
G10L15/00; (IPC1-7): G10L15/10; G10L15/06;
G10L15/22; G10L15/24

- European:

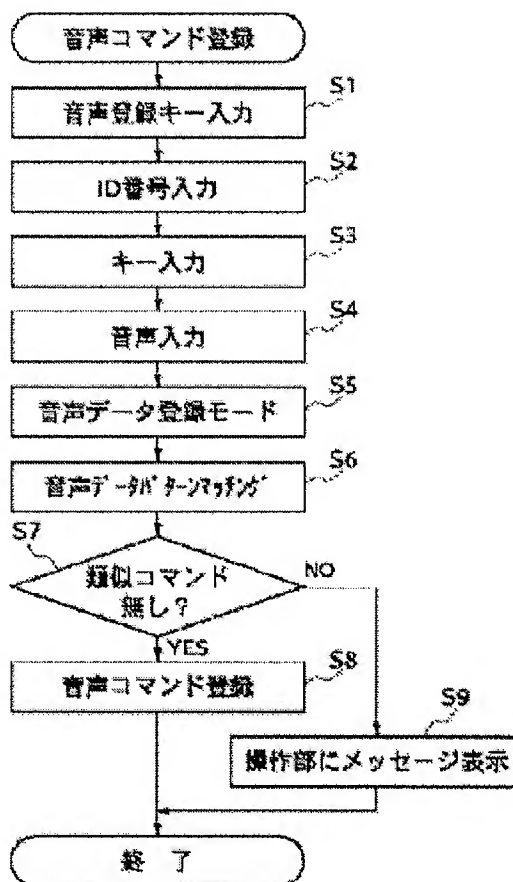
Application number: JP19990064653 19990311

Priority number(s): JP19990064653 19990311

Report a data error here

Abstract of JP2000259172

PROBLEM TO BE SOLVED: To make the device be suitable for the use of many unspecified persons without increasing the size of a storage medium. **SOLUTION:** By pressing down a voice registration key on a control section (S1), an ID number is inputted (S2). Then, while pressing down a desired key, voice data corresponding to the key are inputted (S3 to S4). Then, confirmation is made to check the fact that the key input and the voice data input are performed simultaneously and the device is set to a voice data registration mode (S5). Then, the degree of similarity is computed by conducting a pattern matching of inputted voice data and the voice data registered in a dictionary (S6). If there is no similar command, the inputted voice data are registered in the dictionary (S8). If there exists a similar command, a prescribed message is displayed on the control section (S9).



Data supplied from the esp@cenet database - Worldwide

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2000-259172

(P2000-259172A)

(43) 公開日 平成12年9月22日 (2000.9.22)

(51) Int.Cl. ⁷	識別記号	F I	タームコード [*] (参考)
G 1 0 L 15/10		G 1 0 L 3/00	5 3 1 K 5 D 0 1 J
15/06			5 2 1 V
15/22			5 2 1 N
15/24			5 2 1 J
			5 6 1 C

審査請求 未請求 請求項の数10 O L (全 8 頁) 最終頁に続く

(21) 出願番号 特願平11-64653

(22) 出願日 平成11年3月11日 (1999.3.11)

(71) 出願人 000001007

キヤノン株式会社

東京都大田区下丸子3丁目30番2号

(72) 発明者 鈴木 隆史

東京都大田区下丸子3丁目30番2号 キヤ

ノン株式会社内

(74) 代理人 100081880

弁理士 渡部 敏彦

Fターム (参考) 5D015 AA02 GG03 HH06 HH13 LL05

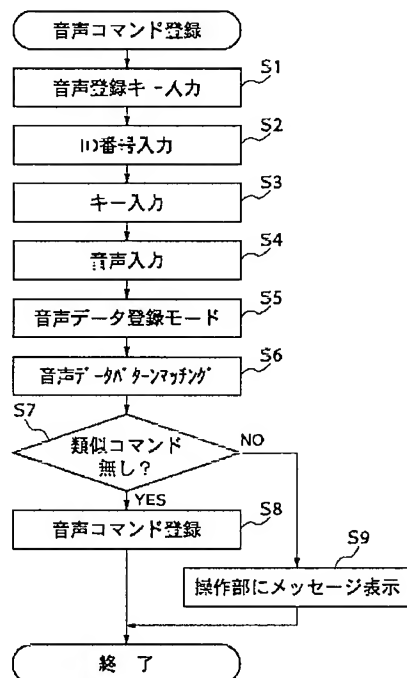
LL08

(54) 【発明の名称】 音声認識装置と音声データの認識方法

(57) 【要約】

【課題】 記憶媒体の大型化を招来することもなく、不特定多数人の使用に好適するようにした

【解決手段】 操作部上の音声登録キーを押下して (S1) ID番号を入力し (S2)、次いで所望のキーを押下しながら該キーに対応する音声データを入力し (S3→S4)、キー入力と音声データ入力が同時になされたことを確認して装置を音声データ登録モードに設定し (S5)、入力音声データと辞書登録されている音声データとをパターンマッチングして類似度を算出する (S6)。そして、類似コマンドがない場合は入力音声データを辞書登録し (S8)、類似コマンドがある場合は所定メッセージを操作部に表示する (S9)。



【特許請求の範囲】

【請求項1】 ユーザの識別情報を入力する識別情報入力手段と、音声データを記憶して蓄積する蓄積手段と、音声データを入力する音声入力手段と、前記識別情報入力手段により入力された識別情報と前記音声入力手段により入力された音声データとを対応付ける対応付け手段と、前記識別情報に対応付けられた入力音声データが蓄積音声データとして前記蓄積手段に既に蓄積されているか否かを判断する判断手段と、該判断手段により前記入力音声データが前記蓄積手段に蓄積されていないと判断されたときは前記入力音声データの前記蓄積手段への新規登録を指示する第1の登録指示手段とを備えていることを特徴とする音声認識装置。

【請求項2】 動作内容を指令する指令情報を入力する入力手段を有すると共に、該入力手段への指令情報の入力操作と同時に前記音声入力手段への音声データの入力が行なわれたときは音声データ登録モードに設定され、前記判断手段は、音声データ登録モードに設定されたときに前記入力音声データと登録音声データとの類似度を算出する類似度算出手段と、該類似度算出手段の算出結果に応じて入力音声データの前記蓄積手段への新規登録の可否を決定する登録可否決定手段とを具備していることを特徴とする請求項1記載の音声認識装置。

【請求項3】 前記判断手段は、前記類似度算出手段により算出された類似度が所定値以上のときは類似音声データの表示指令を発する表示指令手段を有していることを特徴とする請求項2記載の音声認識装置。

【請求項4】 前記識別情報入力手段により識別情報が入力された後に前記音声入力手段による音声データのみが入力されたときは音声データ実行モードに設定され、前記判断手段は、前記音声データ実行モードに設定されたときに入力音声データと登録音声データの類似度を算出する類似度算出手段と、該類似度算出手段の算出結果に応じて入力音声データの指示に基づいた動作処理を実行する動作処理実行手段を具備していることを特徴とする請求項1又は請求項2記載の音声認識装置。

【請求項5】 前記判断手段は、前記類似度算出手段により算出された類似度が所定値以下の場合は入力音声データの蓄積手段への登録指令を発する第2の登録指示手段と、前記類似度算出手段により算出された類似度が所定範囲内の場合は類似音声データの表示指令を発する表示指令手段とを有していることを特徴とする請求項4記載の音声認識装置。

【請求項6】 ユーザの識別情報を入力する識別情報入力ステップと、音声データを入力する音声入力ステップと、前記識別情報入力ステップにより入力された識別情報と前記音声入力ステップにより入力された音声データとを対応付ける対応付けステップと、前記識別情報に対応付けられた入力音声データが蓄積音声データとして蓄積手段に既に蓄積されているか否かを判断する判断ス

tepと、該判断ステップにより前記入力音声データが前記蓄積手段に蓄積されていないと判断されたときは前記入力音声データの前記蓄積手段への新規登録を指示する第1の登録指示ステップとを含んでいることを特徴とする音声データの認識方法。

【請求項7】 動作内容を指令する指令情報の入力操作と同時に前記音声データの入力が行なわれたときは音声データ登録モードに設定し、

前記判断ステップは、音声データ登録モードに設定したときに前記入力音声データと登録音声データとの類似度を算出する類似度算出ステップと、該類似度算出ステップの算出結果に応じて入力音声データの前記蓄積手段への新規登録の可否を決定する登録可否決定ステップとを含んでいることを特徴とする請求項6記載の音声データの認識方法。

【請求項8】 前記判断ステップは、前記類似度算出ステップで算出された類似度が所定値以上のときは類似音声データの表示指令を発する表示指令ステップを含むことを特徴とする請求項7記載の音声データの認識方法。

【請求項9】 前記識別情報入力ステップで識別情報が入力された後に前記音声入力で音声データのみが入力されたときは音声データ実行モードに設定し、

前記判断ステップは、前記音声データ実行モードに設定したときに入力音声データと登録音声データの類似度を算出する類似度算出ステップと、該類似度算出ステップの算出結果に応じて入力音声データの指示に基づいた動作処理を実行する動作処理実行ステップとを含んでいることを特徴とする請求項7又は請求項8記載の音声データの認識方法。

【請求項10】 前記判断ステップは、前記類似度算出ステップで算出された類似度が所定値以下の場合は入力音声データの蓄積ステップへの登録指令を発する第2の登録指示ステップと、前記類似度算出ステップで算出された類似度が所定範囲内の場合は類似音声データの表示指令を発する表示指令ステップを含むことを特徴とする請求項9記載の音声データの認識方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は音声認識装置と音声データの認識方法に関し、より詳しくは、音声により入力された制御コマンドの内容を認識する音声認識装置と音声データの認識方法に関する。

【0002】

【従来の技術】近年、パーソナルコンピュータ（以下、「パソコン」という）やカーナビゲーション（以下、「カーナビ」という）の分野では音声データをコマンドとして入力し、該音声データを認識することにより所望の情報処理を行なうことのできる機種が普及してきている。

【0003】斯かる音声認識は、従来より、声紋や音声

を区切り、音声の高低等をパターンマッチングすることにより行なっている。すなわち、従来では、使用するコマンドと該コマンドを発声したときの音声データとを予め対応付けてメモリ等の記憶媒体に辞書として登録しておき、入力された音声データと登録されているコマンドの音声データの類似度を算出してパターンマッチングを行ない、該類似度の最大値を所望コマンドとして選択し、該選択されたコマンドに基づいて所望の情報処理を行なっている。

【0004】

【発明が解決しようとする課題】しかしながら、上記従来の音声データの認識方法では、音声データを入力して音声の認識を行なうためには、使用する全てのコマンドを予め記憶媒体に辞書として登録しておかなければならず、このため使用し得る全てのコマンドに対してユーザ自身がコマンドを発声し、コマンドを読みながら登録キーと対応させて登録するという操作を必要としていた。

【0005】すなわち、従来の音声データの認識方法では、個人で購入するパソコンやカーナビにおいては、購入した特定人が主として音声認識機能を利用するため、通常は該特定人に関する音声データを登録するのみで対処することができる。

【0006】しかしながら、業務用の複写機やファクシミリ装置、プリンタ、或いはこれらの機能を複合したデジタル複合機等、不特定多数人が使用する機器の場合は、使用する可能性のある多くの人の音声データを記憶媒体に辞書として登録しておく必要があり、したがって大容量の記憶媒体が必要になるという問題点があった。

【0007】また、このような不特定多数人が使用する機器においては、登録されるコマンドについてもユーザによって使用頻度が異なり、したがって特定のユーザにとっては全く使用しないコマンドであっても他のユーザが使用する可能性があるために記憶媒体に斯かるコマンドの音声データを登録しておく必要があり、記憶媒体を効率良く使用することができないという問題点があった。

【0008】このため、上述したデジタル複合機等の不特定多数人が使用する機器では音声認識機能を搭載した機種が未だ存在していないというのが現状である。

【0009】本発明はかかる事情に鑑みてなされたものであり、記憶媒体の大型化を招来することもなく、不特定多数人の使用に好適した音声認識装置と音声データの認識方法を提供することを目的とする。

【0010】

【課題を解決するための手段】上記目的を達成するために本発明に係る音声認識装置は、ユーザの識別情報を入力する識別情報入力手段と、音声データを記憶して蓄積する蓄積手段と、音声データを入力する音声入力手段と、前記識別情報入力手段により入力された識別情報と前記音声入力手段により入力された音声データとを対応

付ける対応付け手段と、前記識別情報に対応付けられた入力音声データが前記蓄積音声データとして前記蓄積手段に既に蓄積されているか否かを判断する判断手段と、該判断手段により前記入力音声データが前記蓄積手段に蓄積されていないと判断されたときは前記入力音声データの該蓄積手段への新規登録を指示する第1の登録指示手段とを備えていることを特徴としている。

【0011】また、本発明に係る音声データの認識方法は、ユーザの識別情報を入力する識別情報入力ステップと、音声データを入力する音声入力ステップと、前記識別情報入力ステップにより入力された識別情報と前記音声入力ステップにより入力された音声データとを対応付ける対応付けステップと、前記識別情報に対応付けられた入力音声データが前記蓄積音声データとして前記蓄積手段に既に蓄積されているか否かを判断する判断ステップと、該判断ステップにより前記入力音声データが前記蓄積手段に蓄積されていないと判断されたときは前記入力音声データの該蓄積手段への新規登録を指示する第1の登録指示ステップとを含んでいることを特徴としている。

【0012】尚、本発明の他の特徴は下記の発明の実施の形態の記載から明らかとなろう。

【0013】

【発明の実施の形態】以下、本発明の実施の形態を図面に基いて詳説する。

【0014】図1は本発明に係る音声認識装置としての複写機の一実施の形態を示すブロック構成図であって、該複写機はコピー動作に関するコマンド情報の入力等を行なう操作部1と、アナログ音声データを入力して該アナログ音声データをデジタル音声データに変換するマイク等からなる音声入力部2と、該音声入力部2からのデジタル音声データに対して所定の音声認識処理を行なう音声認識部3と、原稿画像を読み取ってデジタル画像データに変換するCCD等からなる画像入力部4と、該画像入力部4からのデジタル画像データに対して所定の画像処理を行なうASIC (Application Specific Integrated Circuit) 等のハード回路やソフト処理回路を備えた画像処理部5と、該画像処理部5で画像処理された画像データを出力するプリンタ (レーザービームプリンタ、インクジェットプリンタなど) やモニタ (CRT、LCDなど) 等の画像出力部6と、上記各構成要素に接続されてこれら各構成要素を制御するドライバ部7とから構成されている。

【0015】また、音声認識部3は、図2に示すように、音声入力部2からのデジタル音声データが入力される音声データ入力部8と、操作部1からのコマンド情報を入力する操作部情報入力部9と、操作部情報とデジタル音声データとを対応させて記憶するRAMやハードディスク等の記憶媒体で構成された辞書データ蓄積部10と、音声データ入力部8からのデジタル音声データと辞

書データ蓄積部10に蓄積された辞書音声データとの間でパターンマッチングを行ない、各コマンド相互間の類似度を算出するパターンマッチング部11と、ドライバ部7からの各種コマンド情報等が入力されるドライバ情報入力部12と、前記パターンマッチング部11、前記操作部情報入力部9及び前記ドライバ情報入力部12からの情報に基づいて所定のコマンド処理を行なうコマンド処理部13と、コマンド処理部13から出力されたコマンド情報を適宜ドライバ部7や操作部1に送信するコマンド出力部14とを備えている。

【0016】図3は操作部1の平面図であって、該操作部1は、各種キー群15とモード表示部16とを有している。

【0017】各種キー群15は、具体的には、数字キーやIDキー17a等を備えたテンキー17と、コピー動作を実行するときに操作するコピーキー18と、コピー動作を中断するときに操作するストップキー19と、コマンドの音声データを登録するときに操作する音声登録キー20と、リセットキー21とを有している。また、モード表示部16の上方部適所には液晶表示パネル16aが設けられている。

【0018】図4は音声コマンドの登録手順を示すフローチャートである。

【0019】まず、ステップS1では音声登録キー20を押下する。これにより、操作部1の液晶パネル表示部16aには、図5に示すように、「ID番号をテンキーで入力して下さい。」のメッセージが表示される。

【0020】次に、ステップS2ではテンキー17を操作して所定のID番号（識別情報）を入力し、次いで各種キー群15のうちから選択された1個のキーを押下し（ステップS3）、該キーを押下しながら押下したキーの「読み」を音声データとして入力する（ステップS4）。例えば、コピーキー18を押下した場合は「コピー」と発声して音声入力部2に該音声データ「コピー」を入力し、テンキー17の中の「1」を押下した場合は「いちまい」と発声して音声入力部2に該音声データ「いちまい」を入力する。そして、キー入力と音声入力とが同時に行なわれていると判断されると、続くステップS5で本複写機は音声データ登録モードに設定され、次いでステップS6に進み、パターンマッチング部11による音声データのパターンマッチングを行ない、入力されたコマンドと辞書データ蓄積部10に登録されている同一コマンドの類似度を算出する。

【0021】次に、ステップS7に進んで類似コマンドが辞書データ蓄積部10に登録されていないか否かを判断する。ここで、類似度の算出結果により判断され、本実施の形態では入力コマンドと辞書データ蓄積部10に登録されている登録コマンドとが完全一致する場合を類似度「100」とし、類似度が「90」以上の場合は入力コマンド

と登録コマンドとが略同一と認められると判断し、類似度が「80以上90未満」の場合は入力コマンドに対して候補となり得る候補コマンドが登録されていると判断し、類似度が「80」未満の場合は未登録コマンドが入力されたと判断し、これにより類似コマンドの既登録か否かを判断する。例えば、辞書音声データとして「いちまい」、「はちまい」、「さんまい」、「こびー」が辞書データ蓄積部10に登録されている場合に、音声データ「いちまい」が入力された場合は、辞書音声データ「いちまい」との間の類似度は「100」と判断され、辞書音声データ「はちまい」との間の類似度は「85」と判断され、辞書音声データ「さんまい」との間の類似度は「30」と判断され、辞書音声データ「コピー」との間の類似度は「5」と判断され、類似コマンドが80以上の場合に類似コマンドが登録されていると判断する。

【0022】そして、類似コマンドが登録されていないと判断された場合、すなわちステップS7の答が肯定（Yes）の場合は入力された音声データをコマンド情報として辞書データ蓄積部10に登録し（ステップS8）、音声コマンドの登録処理を終了する。

【0023】一方、ステップS7の答が否定（No）、例えば、「いちまい」という音声データ入力を行なった場合、パターンマッチング部11でのマッチング結果が類似度「90」となった場合は類似コマンドが存在すると判断してステップS9に進み、図6に示すように、液晶表示パネル16aに「今のコマンドは「1枚」ですか？」のメッセージを表示すると共に「YES」、「NO」の選択キー16bを表示する（ステップS9）。そして、「YES」の場合はYESキーを押下し、ユーザは既に「いちまい」が登録済みであることを確認し、音声コマンドの登録処理を終了する。

【0024】尚、前記液晶表示パネル16aで「NO」キーを選択したときはステップS1から登録手順をやり直すこととなる。

【0025】図7は音声コマンド認識手順のフローチャートである。

【0026】ステップS11でユーザが操作部1のIDキー17aを押下すると、上述した図5と同様、操作部1の液晶パネル表示部16aには、「ID番号をテンキーで入力して下さい。」のメッセージが表示される。

【0027】次に、ステップS12ではテンキー17を操作して所定のID番号を入力し、音声データを音声入力部2に入力し（ステップS13）、音声データ実行モードにモード設定する（ステップS14）。

【0028】次いで、ステップS15に進み、パターンマッチング部11で音声データのパターンマッチングを行ない、入力コマンドと登録コマンドとの類似度を算出する。

【0029】続くステップS16では類似コマンドが登

録されているか否かを判断し、類似コマンドが登録されている場合はステップS17でその類似コマンドが複数登録されているか否かを判断する。そして、その答が否定(N o)、すなわち類似コマンドが1個の場合は入力された音声コマンドに対応した所望のコピー動作を実行し(ステップS18)、処理を終了する。

【0030】また、ステップS17の答が肯定(Y e s)、すなわち類似コマンドが複数検索された場合は、操作部1の液晶表示パネル16aに所定のメッセージを表示する。例えば、音声データ「いちまい」が入力された場合、「はちまい」との類似度は「85」であるため、「いちまい」と「はちまい」という2つの候補コマンドが存在することとなり、図8に示すように、液晶表示パネル16aには「今のコマンドはどちらですか?」のメッセージを表示すると共に2つの候補コマンド16c、すなわち「1枚」及び「8枚」という候補コマンドを液晶表示パネル16aに表示する。そして、ステップS20で音声コマンド、例えば候補コマンド「1枚」キーを押して音声コマンド「いちまい」を選択し、次いで、ステップS18に進み、ドライバ部7は、斯かるコマンドに基づいてコピー処理を行ない処理を終了する。

【0031】また、ステップS16で類似コマンドがないと判断されたとき、すなわち未登録の音声データが入力されたときは、ステップS21に進み、図9に示すように、液晶表示パネル16aに「未登録です。登録するキーを押して下さい。」のメッセージを表示する。

【0032】次いで、ステップS22で各種キー群15の中から所望のキーを選択して操作し所望の音声コマンドを辞書データ蓄積部10に登録した後、ステップS18に進んでドライバ部7は、斯かるコマンドに基づいてコピー処理を行ない処理を終了する。

【0033】このように本実施の形態によれば、各個人のID番号と音声コマンドとを対応付けて所望の音声データをその都度登録するようにしているので、使用し得るコマンドの全てについて予め登録しておくという手間が省けると共に、各個人の各々が頻繁に使用する音声コマンドのみを各自の判断で任意に登録することができ、

記憶媒体の容量低減化を図ることができる。

【0034】

【発明の効果】以上詳述したように本発明によれば、ユーザの識別情報と対応付けて必要に応じて所望の音声データを登録しているので、各自の使用頻度に応じてユーザが必要と考える音声データのみを登録すればよく、業務用の複写機等、不特定多数人が使用する機種に対しても比較的容量の小さい記憶媒体であっても対処することが可能となる。

【図面の簡単な説明】

【図1】本発明に係る音声認識装置としての複写機の一実施の形態を示すブロック構成図である。

【図2】音声認識部の詳細を示すブロック構成図である。

【図3】操作部の詳細を示す平面図である。

【図4】音声コマンドの登録手順を示すフローチャートである。

【図5】音声コマンドの登録時における操作部の一例を示す平面図である。

【図6】音声コマンドの登録時における操作部の他の例を示す要部平面図である。

【図7】音声コマンドの認識手順を示すフローチャートである。

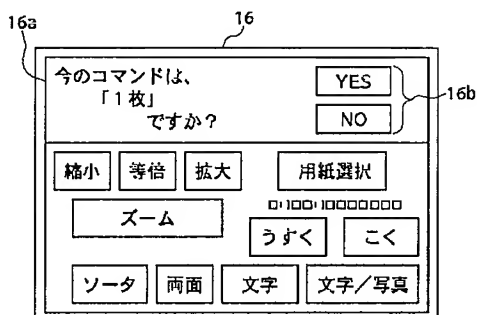
【図8】音声コマンドの認識時における操作部の一例を示す要部平面図である。

【図9】音声コマンドの認識時における操作部の他の例を示す要部平面図である。

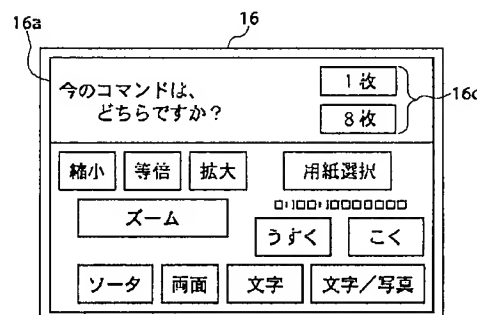
【符号の説明】

- 1 操作部(入力手段)
- 4 音声入力部(音声入力手段)
- 10 辞書データ登録部(蓄積手段)
- 11 パターンマッチング部(類似度算出手段)
- 13 コマンド処理部(対応付け手段、登録可否決定手段、動作処理実行手段)
- 14 コマンド出力部(第1及び第2の登録指示手段、表示指令手段)
- 17a IDキー(識別情報入力手段)

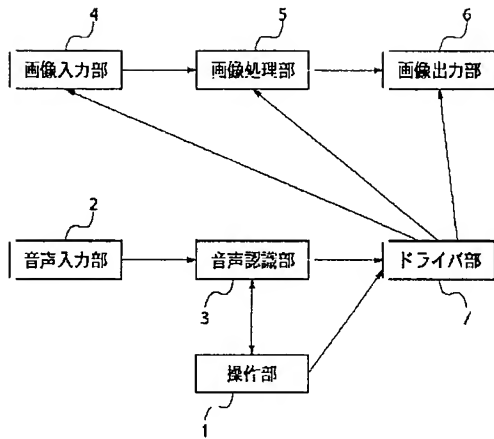
【図6】



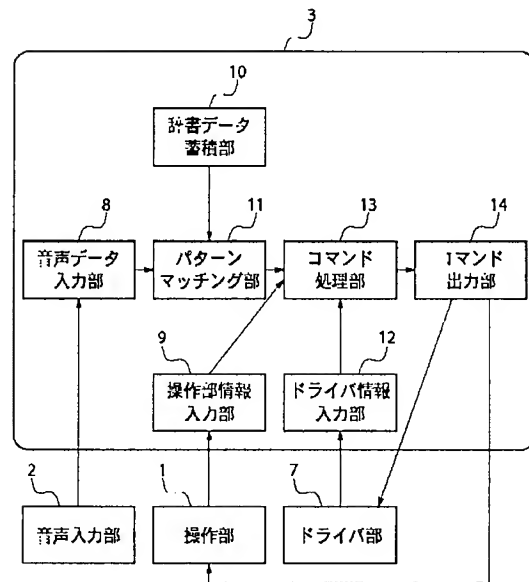
【図8】



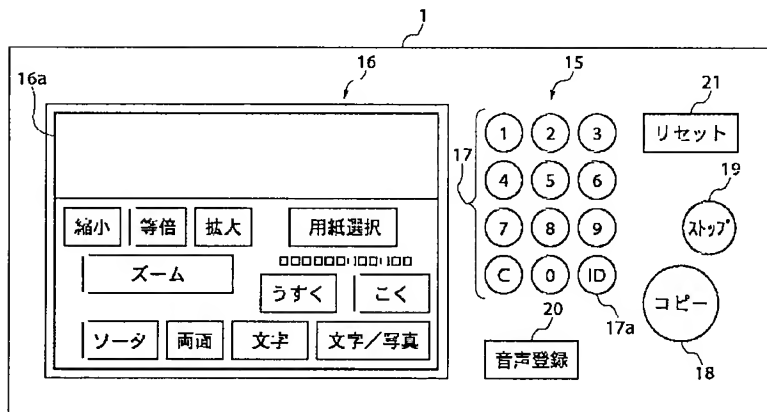
【図1】



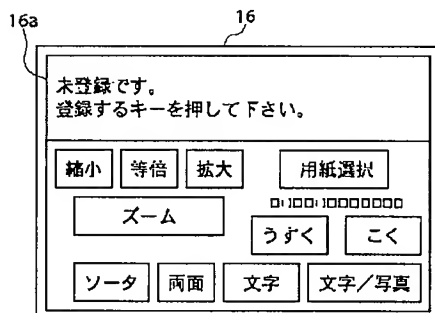
【図2】



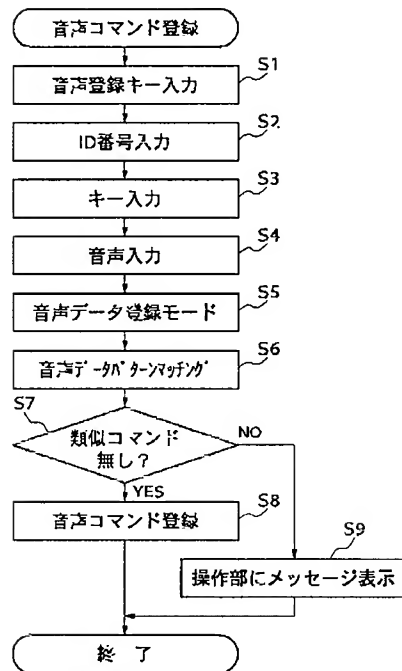
【図3】



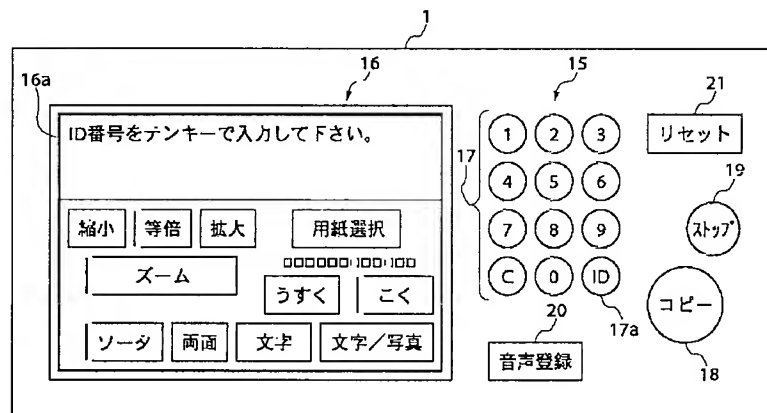
【図9】



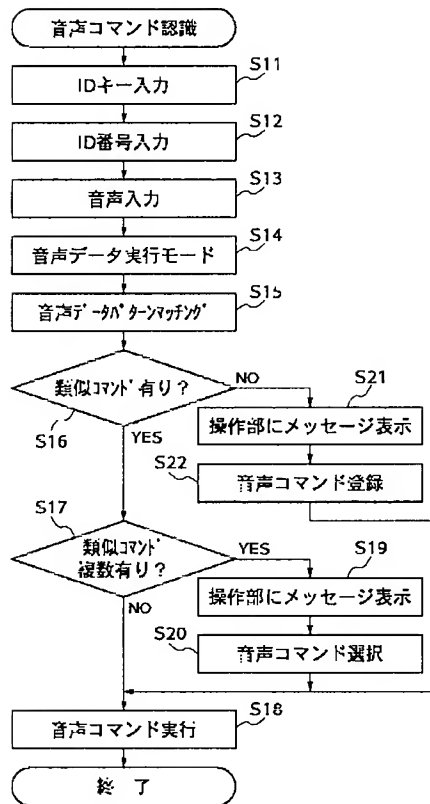
【図4】



【図5】



【図7】



フロントページの続き

(51)Int.Cl.⁷

識別記号

F I
G 1 0 L 3/00

5 7 1 G

(参考)